



Viewpoints

February 2007

Knowledge-Management Tools

Technology Analyst: Marcelo Hoffmann

+1 650 859 3680; fax: +1 650 859 4544; electronic mail: mhoffmann@srbc-bi.com

Ontologies

Why is this topic significant?

The **SEMANTIC WEB** in **Areas to Monitor** in the Technology Map describes an attempt by the World Wide Web Consortium to structure and further integrate the Web, thus making the Web more amenable to automated processing and understanding by computers without human intermediation. Ontologies support the evolution of the Semantic Web with, among other elements, widespread integration of complex online knowledge stores and large-scale databases. This Viewpoints describes the significance of ontologies and describes an effort to integrate multiple domain-specific ontologies under so-called upper ontologies.

Ontology is becoming ever more important for the scientific and commercial development of smart computer systems. An ontology defines the terms to describe and represent an area of knowledge. People, databases, and software applications that need to share domain information use ontologies. Ontologies typically include computer-usable definitions of basic concepts in the domain at hand and the relationships among these concepts. Ontologies encode knowledge in a domain and also knowledge that spans domains. In this way, ontologies make knowledge reusable.

According to Tom Gruber, a computer scientist and researcher associated with Stanford University's Knowledge Systems Laboratory, ontology is "a specification of a conceptualization." For other computer scientists, an ontology is typically a hierarchical data structure that contains all



An SRI International Business Partner

Viewpoints, a monthly bulletin, alerts members to commercially significant developments in technology areas selected by their company. See Explorer's Technology Maps for a comprehensive assessment of issues, uncertainties, and opportunities. Membership includes access to Explorer's Inquiry Service. To use the service, contact the analyst directly, or e-mail inquiry@srbc-bi.com.

the relevant entities and their relationships and rules within that domain.

Ontologies generally include descriptions or definitions of:

- *Individuals*. Individuals are the most basic components of an ontology. The individuals in an ontology may include concrete objects such as people, animals, vehicles, molecules, and planets, as well as abstract concepts such as numbers and words. One of the general purposes of an ontology is to provide a means of classifying individuals.
- *Classes*. This category includes sets, collections, or types of objects and may contain within them, individuals, other classes, or a combination of them. Examples of classes include *molecule* as the class of all molecules, *vehicle* as the class of all vehicles, and *flower* as the class of all flowers.
- *Attributes*. Attributes are properties, characteristics, or parameters that objects can have, and sometimes share among each other. Typically, each attribute has at least a name and value; a “Mercedes Benz sedan” object will have attributes such as:
 - Name: Mercedes Benz 500
 - Number of doors: 4
 - Engine size: 5-liter
 - Transmission: 6-speed automatic
- *Relationships*. Relationships include ways in which objects can be related to one another. For the Mercedes Benz 500 object, relationships may include where the particular product line fits among various others in the Mercedes Benz product line (by size, price, engine size, and the like). Researchers often describe relationships in treelike structure that depicts where one object fits with respect to another, or as a “meronymy relation” (a description that ties parts of an object or class; for example, Mercedes Benz 500 is-example-of-automotive-sedan).

Computer scientists, artificial-intelligence researchers, and software developers specify ontologies to enable knowledge sharing and reuse across applications and domains of expertise. Thus, an ontology can become a formal description of the concepts and relationships that exist and are significant for a human or software agent or a community of agents. In practical terms, by using commonly shared ontologies, software developers can make queries and assertions about the subject matter covered by the ontology consistent with the theory specified by a

specific ontology. And ideally, computers can use ontologies to replicate inferencing done by humans, using some degree of artificial intelligence.

Ontologies are usually expressed in a logic-based computer language, enabling automated understanding about the classes, properties, and relations within them. The more sophisticated ontology tools can perform automated reasoning and enable intelligent applications such as conceptual or semantic search and retrieval, software agents, decision support, speech and natural-language understanding, knowledge management, intelligent databases, and electronic commerce.

Ontologies figure prominently in the development of complex knowledge-based systems and in the emerging semantic Web as a way to represent the semantics of documents and enable the semantics to be used by Web applications and intelligent agents. Ontologies are particularly useful as a way of structuring and defining the meaning of metadata terms.

The term *ontology* has historically described artifacts with different degrees of structure that range from simple taxonomies (such as the Yahoo hierarchy for search terms) to metadata schemes (such as the Dublin Core Metadata Initiative, an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models; <http://dublincore.org/>) to logical theories. Dublin Core is an “upper ontology,” because it models common objects that are generally useful across a wider range of domain-specific ontologies—in contrast to narrower ontologies that would individually describe a particular card game, such as poker, blackjack, or bridge.

Because domain-specific ontologies represent concepts in narrowly defined ways, they are often incompatible and inconsistent with each other. As more complex knowledge-management and knowledge-based systems that rely on domain-specific ontologies expand, they often need to merge domain ontologies into a more general representation (for example, to integrate various card-game ontologies into a larger ontology that represents games more generally). Different domain-specific ontologies can also develop independently because of different perceptions of the domain based on cultural background and education of the developers, or because a different representation language was the choice. Both cases present challenge to ontology developers and maintainers. Currently, merging ontologies requires a labor-intensive effort that is time-consuming and rather expensive. Although using an

upper ontology to provide a common definition of core terms can make this process manageable, the area is not well developed and still under research, but with ongoing efforts to make integration easier.

A personal conversation with Peter Yim, the coordinator and host of Ontolog Forum (<http://ontolog.cim3.net/>), an online community of practice in ontologies, brought out a potentially significant development for the interoperability of disparate ontologies. On 15 March 2006, a group of computer scientists convened the Upper Ontology Summit, a specialized workshop in conjunction with the U.S. National Institute of Standards and Technology (Gaithersburg, Maryland) and its Interoperability Week (<http://www.mel.nist.gov/div826/msid/sima/interopweek/index.htm>). The purpose of the Upper Ontology Summit workshop (<http://ontolog.cim3.net/cgi-bin/wiki.pl?UpperOntologySummit>), according to the organizers, was “to bring together those who recognize the value of open upper (high-level) ontologies, and in particular, the developers or maintainers of the public versions of existing upper ontologies to find a way to interrelate those ontologies....” In particular, the conveners wanted to “develop the mechanism and resources to relate existing upper ontologies to each other in a manner that will increase reuse of knowledge among them, and thereby facilitate semantic interoperability among other ontologies that are linked to them.”

Convening participants suggested that to achieve interoperability they would seek to create a simplified overarching upper ontology that included a compatible subset of all the linked upper ontologies. Although a number of high-level ontologies are now available for specific industrial and scientific domains, most are designed with a relatively narrow perspective, have a limited set of users, lack the means to interoperate across ontologies, and thus have limited value. Ideally, the new, interoperable ontology would have enough expressive power for practical applications and to synergize with specialized upper ontologies.

CURRENTLY AVAILABLE TECHNOLOGY AREAS

Advanced Silicon Microelectronics/ULSI	Nanobiotechnology
Biocatalysis	Nanoelectronics
Biomaterials	Nanomaterials
Biopolymers	Novel Ceramic/Metallic Materials
Biosensors	Optoelectronics/Photonics
Connected Cars	Organic Electronics
Connected Homes	Pervasive Computing
Engineering Polymers	Photovoltaics
Fiber-Optic Sensors	Polymer-Matrix Composites
Flat-Panel Displays	Portable Intelligence
Fuel Cells	Portable Power
Genomics	Renewable Energy Technologies
Knowledge-Based Systems	RFID Technologies
Knowledge-Management Tools	Robotics
Membrane Separation	Smart Materials
MEMS/Micromachining	Solid-State Microsensors
Mobile Communications	Virtual Environments

Explorer provides up-to-date evaluations and analysis of business opportunities arising from technology commercialization. Explorer filters information, provides perspectives on developments, and allows users to focus on what they need to help their decision processes. For more detailed information about Explorer, visit our Web site at www.sric-bi.com; e-mail: info@sric-bi.com; or phone +1 650 859 4600 in the United States; +44 (0) 20 8686 5555 in Europe; +81 3 3222 6501 in Asia.